JOURNAL OF MULTIVARIATE ANALYSIS 43, 147–170 (1992)

Prediction and Fundamental Moving Averages for Discrete Multidimensional Harmonizable Processes

MARC H. MEHLMAN

University of Pittsburgh at Johnstown

Communicated by the Editors

The moving average representations of discrete multidimensional stationary processes are generalized to fundamental moving average representations of weakly harmonizable processes. For strongly harmonizable processes, necessary and sufficient conditions on covariance functions are obtained for the existence of such moving average representations. These are used in obtaining least squares prediction formulae for such processes. © 1992 Academic Press, Inc.

1. Introduction

The purpose of this paper is to discuss and generalize the concept of fundamental moving average representation from multidimensional discrete stationary processes to weakly harmonizable processes and to employ these representations in prediction theory. Weakly harmonizable processes are a proper subset of all bounded continuous processes and they are a natural extension of stationary processes. This class of processes, whose study is amenable to Fourier analytic methods, is of interest for applications such as prediction and filtering problems, among others.

In Section 2 a brief account of the spectral representation of discrete stationary and harmonizable processes that is utilized for the main results is given (see [7], where the continuous parameter case is considered too). The notion of a virile moving average representation of a harmonizable process is recalled and a concept of rank is discussed in Section 2.3 for later use. Also included here are some extensions to the clasical results obtained by Y. Rozanov [10] for stationary processes. The main results of this paper on extrapolation and moving average representations are established in Sections 3 and 4.

Received June 11, 1991; revised February 13, 1992.

AMS subject classification: 60.

Key words and phrases: Harmonizable processes, moving averages, prediction.

0047-259X/92 $5.00

2. Harmonizable Processes

In the following work there is always a probability space, (Ω, Σ, P), in the background, even if it is not always explicitly mentioned.

2.1. *Definitions*

Definition 2.1. For $p \geqslant 1$ define $L_0^p(P)$ to be all complex valued $f \in L^p(P)$ such that $E(f) = 0$, where $E(f) \stackrel{\text{def}}{=} \int_\Omega f(\lambda)\, P(d\lambda)$ is the expectation of f.

Letting $\mathscr{M}_{n,m}$ denote the set of $n \times m$ complex valued matrices, it is frequently advantageous to view $[L_0^2(P)]^n$ as a left $\mathscr{M}_{n,n}$ module: if $X, Y \in [L_0^2(P)]^n$ and $A, B \in \mathscr{M}_{n,n}$ then $AX + BY \in [L_0^2(P)]^n$. There exists an $\mathscr{M}_{n,n}$-valued Gramian (see [5]) defined on $[L_0^2(P)]^n$ given by $[X, Y] \stackrel{\text{def}}{=} E(XY^*) \in \mathscr{M}_{n,n}$. (Here $*$ denotes the adjoint operator, i.e., the conjugate transpose operator.) In fact, $[L_0^2(P)]^n$ may be viewed as a Hilbert space with $(X, T) \stackrel{\text{def}}{=} \operatorname{tr}[X, Y]$ and $\|X\|_{[L_0^2(P)]^n} \stackrel{\text{def}}{=} \operatorname{tr}[X, X]$.

We will be considering second order discrete random processes (those with $\mathbf{Z}$ for their index sets). Associated with the topological group, $\mathbf{Z}$, is its dual group: $\hat{\mathbf{Z}}$ is the unit circle in C, denoted by T. As usual, T will be thought of as the interval $[-\pi, \pi)$. $\mathscr{B}$ will denote the set of Borel subsets of T.

Definition 2.2. Let $X_t = (X_t^{(1)}, \ldots, X_t^{(n)})^T$ be an n-dimensional random process. Define

$$H_X^-(t) \stackrel{\text{def}}{=} \begin{cases} \overline{\mathrm{sp}}\{X_s^{(j)} : s \leqslant t, 1 \leqslant j \leqslant n\} & \text{if } t \in \mathbf{Z}, \\ \overline{\mathrm{sp}}\{X_s^{(j)} : s \in \mathbf{Z}, 1 \leqslant j \leqslant n\} & \text{if } t = \infty, \\ \bigcap_{s \in \mathbf{Z}} H_X^-(s) & \text{if } t = -\infty, \end{cases}$$

where closure is taken in $L_0^2(P)$. The space $H_X^-(\infty)$ is referred to as the *space of observables* of X_t. Given an $[L_0^2(P)]^n$-valued vector measure, $Z(\cdot)$, on $(T, \mathscr{B})$, for every $\Delta \in \mathscr{B}$ let

$$H_Z^-(\Delta) \stackrel{\text{def}}{=} \overline{\mathrm{sp}}\{Z^{(j)}(\Delta' \cap \Delta) : 1 \leqslant j \leqslant n, \Delta' \in \mathscr{B}\},$$

where closure is again taken in $L_0^2(P)$. The notation $l.i.m._{n \uparrow \infty} Y_n = Y$ is used for convergence in mean-square, i.e., that $\lim_{n \uparrow \infty} \|Y_n - Y\|_2 = 0$.

Definition 2.3. A complex valued random process, X_t, is *stationary* (stationary in the wide or Khinchine sense) iff (= if and only if) the covariance function, $r_X(s, t) \stackrel{\text{def}}{=} E(X_s X_t^*)$ of X_t, is continuous and is a function of the difference of its arguments, i.e., if $r_X(s, t) = r_X(s+u, t+u)$ for all u, s, $t \in \mathbf{Z}$. Hereafter $\tilde{r}(t) \stackrel{\text{def}}{=} r(0, t)$.

An equivalent definition, by the classical Bochner Theorem [10], of a stationary process is one whose covariance function can be represented as

$$\tilde{r}(s) = \int_T e^{i\lambda s} F(d\lambda),$$

for a unique non-negative bounded measure $F(\cdot)$ on $(T, \mathscr{B})$.

DEFINITION 2.4. A random process, X_t, taking values in $L_0^2(P)$ is *weakly harmonizable* iff its covariance function can be expressed as

$$r(s, t) = \iint_{T \times T} e^{i\lambda s - i\lambda' t} F(d\lambda, d\lambda'), \tag{2.1}$$

where $F(d\lambda, d\lambda')$ is a positive semi-definite bimeasure on $\mathscr{B} \times \mathscr{B}$, hence of finite Fréchet variation. The above integral is a strict Morse–Transue integral [1]. A random process, X_t, is *strongly harmonizable* iff the bimeasure $F(d\lambda, d\lambda')$ in (2.1) extends to a complex measure (hence has bounded variation in Vitali's sense) on the Borel σ-algebra of $T \times T$. In either case, $F(d\lambda, d\lambda')$ is called the *spectral bimeasure* (or spectral measure when $F(d\lambda, d\lambda')$ is a measure) of X_t.

DEFINITION 2.5. An n-dimensional vector of processes,

$$X_t(\cdot) \stackrel{\text{def}}{=} (X_t^{(1)}(\cdot), ..., X_t^{(n)}(\cdot))^T,$$

is an n-dimensional weakly (strongly) harmonizable or stationary process iff for every $1 \times n$ vector of complex numbers, w, the process $w \cdot X_t$ is weakly (strongly) harmonizable or stationary.

A standard calculation reveals that equivalent to the above definition of an n-dimensional harmonizable random process is to require that its covariance function be representable as

$$\iint_{T \times T} e^{i\lambda s - i\lambda' t} F(d\lambda, d\lambda'),$$

where $F(d\lambda, d\lambda')$ is an $n \times n$ matrix array of bimeasures. Likewise, an equivalent definition for an n-dimensional stationary process is that its covariance function can be represented as $\int_T e^{i\lambda(t-s)} F(d\lambda)$ where $F(d\lambda)$ is an $n \times n$ matrix array of complex valued measures ($F(\Delta)$ will necessarily be a positive semi-definite matrix for all $\Delta \in \mathscr{B}$).

DEFINITION 2.6. An n-dimensional harmonizable process, X_t, has *rank* p iff its spectral bimeasure takes values in the space of $n \times n$ matrices of rank p together with 0_n. If $n = p$ the process is said to have *maximal rank*.

One should note that the rank of the covariance function is different from that of the spectral bimeasure. Rank is not defined for all harmonizable processes. However, when it is defined it is an upper bound for rank $r(s, t)$.

DEFINITION 2.7. An $[L_0^2(P)]^n$-valued measure, $Z(\cdot)$, has *orthogonal increments* (or is said to be *orthogonally scattered*) iff $\Delta \cap \Delta' = \varnothing$ implies that

$$E(Z(\Delta)\, Z^*(\Delta')) = 0_n.$$

2.2. *Spectral Representation*

The following known result gives a characterization of weakly harmonizable processes (see [9]) and will be used below.

THEOREM 2.8. *An n-dimensional process, X_t, is weakly harmonizable iff it has a spectral representation*

$$X_t = \int_T e^{it\lambda} Z(d\lambda), \tag{2.2}$$

where $Z(d\lambda)$ is an $[L_0^2(P)]^n$-valued measure. The process, X_t, is a stationary process iff $Z(d\lambda)$ has orthogonal increments. [The integral in (2.2) is in the Dunford–Schwartz sense.]

The spectral bimeasure, $F(\cdot, \cdot)$, of an n-dimensional harmonizable random process satisfies $F(A, B) = E(Z(A)\, Z^*(B))$. In the stationary case, $Z(d\lambda)$ is orthogonally scattered so that the spectral measure is concentrated on the diagonal of $T \times T$ and can be written as $F(d\lambda)$. This case gives the well known representation of stationary processes due to H. Cramér and A. Kolmogorov.

DEFINITION 2.9. Let X_t be an n-dimensional harmonizable random process with spectral representation $X_t = \int_T e^{it\lambda} Z(d\lambda)$. For $p \in \mathbf{Z}^+$ define an equivalence relation on the set of $p \times n$ matrix valued functions by $A(\cdot) \sim B(\cdot)$ iff $\|\int_T (A - B)(\lambda)\, Z(d\lambda)\|_{[L_0^2(P)]^p} = 0$. Let $L^2(F, p)$, the *spectral domain* of X_t, be the set of equivalence classes $[A(\cdot)]$ such that $\int_T A(\lambda)\, Z(d\lambda) \in [L_0^2(P)]^p$.

If $F(\cdot, \cdot)$ is the spectral measure of X_t and $A(\cdot)$ and $B(\cdot)$ are $p \times n$ matrix valued functions, then $A \sim B$ iff

$$\iint_{T \times T} (A - B)(\lambda)\, F(d\lambda, d\lambda')(A - B)^* (\lambda') = 0_p,$$

where the integrals are defined componentwise. The spectral domain of X_t depends only on the spectral measure $F(\cdot, \cdot)$ of X_t and not on X_t itself since $Y_A \overset{\text{def}}{=} \int_T A(\lambda)\, Z(d\lambda) \in [L_0^2(P)]^p$ iff

$$E(Y_A Y_A^*) = \iint_{T \times T} A(\lambda)\, F(d\lambda, d\lambda')\, A^*(\lambda') \in \mathcal{M}_{p,p}.$$

PROPOSITION 2.10. *Given an n-dimensional strongly harmonizable process, $X_t = \int_T e^{it\lambda}\, Z(d\lambda)$, one has $H_X^-(\infty) = H_Z^-(T)$.*

DEFINITION 2.11. Given a probability space (Ω, Σ, P) and taking any other probability space (Ω', Σ', P'), one can "enlarge" (Ω, Σ, P) to an *augmented probability space*, $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$, by letting $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P}) \overset{\text{def}}{=} (\Omega \times \Omega', \Sigma \times \Sigma', P \otimes P')$.[1]

Let (Ω, Σ, P) be a probability space and $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$ be an augmentation of that probability space. For each $\tilde{\omega} \in \tilde{\Omega}$ one can write $\tilde{\omega} = (\omega, \omega')$ where $\omega \in \Omega$ and $\omega' \in \Omega'$. Given a random process X_t on (Ω, Σ, P) one can identify X_t with a random process $\tilde{X}_t$ on the augmented probability space by letting $\tilde{X}_t(\tilde{\omega}) \overset{\text{def}}{=} X_t(\omega)$. Since the distributions of X_t and $\tilde{X}_t$ are the same, the two random variables are indistinguishable from a probabilistic point of view.

The following theorem, proved by M. M. Rao [9, Theorem 6.1], implicitly uses this identification. The proof involves using the results of Theorem 2.8 along with a Grothendieck type inequality. Clearly, an n-dimensional version also holds.

THEOREM 2.12 (Dilation Theorem). *A random process, X_t, is a weakly harmonizable process iff it has a stationary dilation (Y_t, π), i.e., there exists a stationary process, Y_t, on an augmented probability space $(\tilde{\Omega}, \tilde{\Sigma}, \tilde{P})$ along with an orthogonal projection, $\pi: L_0^2(\tilde{P}) \to L_0^2(P)$, (where $L_0^2(P)$ is considered embedded in $L_0^2(\tilde{P})$) such that $X_t = \pi Y_t$.*

Given the dilation theorem one can immediately obtain Theorem 2.8. However, at the moment, no independent proof of Theorem 2.12 is known for obtaining the representation Theorem 2.8.

[1] For every $A \in \Sigma$ and $A' \in \Sigma'$, $P \otimes P'$ is defined by $P \otimes P'(A \times A') \overset{\text{def}}{=} P(A)\, P'(A')$.

DEFINITION 2.13. An n-dimensional random process is called *splitting* (or a splitting process) iff its covariance function factors as $r(s,t) = G(s)\,G^*(t)$, where $G(\cdot)$ is an $n \times q$ matrix valued function on T.

A random process, X_t, is a splitting process iff $\dim H_X^-(\infty) < \infty$ (see [7, Theorem 4.2]). Furthermore it is shown that weakly harmonizable splitting processes are strongly harmonizable (see [7, Theorem 4.6]). Using this fact and the above dilation theorem, the following lemma was also established.

LEMMA 2.14. *Given an n-dimensional weakly harmonizable process, X_t, and a stationary dilation (Y_t, π), then X_t is strongly harmonizable if $H_X^-(\infty)$ is finite dimensional or has finite codimension in $H_Y^-(\infty)$.*

2.3. *Generalized Moving Averages*

We now generalize the definition of moving averages from that what is commonly used in the literature for stationary processes (the latter moving averages will henceforth be called orthonormal moving averages).

Letting $\delta_o(x) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise} \end{cases}$, one has:

DEFINITION 2.15. A moving average representation of an n-dimensional random process, X_t, is a representation

$$X_t = \sum_{j \in \mathbf{Z}} \hat{c}(j-t)\,\xi_j, \tag{2.3}$$

where

1. $\hat{c}(\cdot)$ is the Fourier transform (taken component-wise) of a function $c: T \to \mathcal{M}_{n,m}$, whose every component is in $L^2(d\lambda)$ and

2. $r_\xi(s,t) = \rho(s,t)\,I_m$ where $\rho(\cdot,\cdot)$ is the covariance function of a one dimensional process.

Furthermore, if

(a) $E(\xi_s \xi_t^*) = 0_m$ when $s \neq t$,

(b) $E(\xi_s \xi_t^*) = \delta_0(t-s)\,I_m$,

(c) ξ_j is a stationary process,

(d) ξ_j is a strongly harmonizable process,

(e) ξ_j is a weakly harmonizable process, then the moving average representation in (2.3) is, respectively, termed

(a) *orthogonal moving average,*
(b) *orthogonal moving average,*
(c) *stationary moving average,*
(d) *strongly harmonizable moving average,*
(e) *weakly harmonizable moving average.*

A random process can have more than one moving average representation. Lemma 2.20 will relate a moving average representation's type (stationary, strongly harmonizable, or weakly harmonizable) with the type of the process it represents.

"Stationary orthonormal moving average" is redundant since every orthonormal moving average must be a stationary moving average. If X_t has an orthonormal moving average representation (2.3), then X_t is a stationary process (see Lemma 2.20 below). Furthermore, "orthogonal stationary moving averages" are equivalent (modulo a multiplicative constant) to orthonormal moving averages.

DEFINITION 2.16. A moving average representation (2.3) has *rank p* iff $c(\lambda)$ has rank p for every $\lambda \in T$. A moving average representation has *full rank m* iff it has rank m.

For weakly (and strongly) harmonizable processes, the following definition is presented:

DEFINITION 2.17. A weakly (strongly) harmonizable moving average (2.3) is a *virile moving average* iff ξ_t is weakly (strongly) harmonizable with spectral measure $\mu_\xi(d\lambda, d\lambda')\, I_m$ and for $N \in \mathbf{Z}^+$, letting

$$c_N(\lambda) \overset{\text{def}}{=} \sum_{|j|>N} \hat{c}(j)\, e^{ij\lambda}$$

one has

$$\lim_{N \uparrow \infty} \iint_{T \times T} c_N(\lambda)\, c_N^*(\lambda')\, \mu_\xi(d\lambda, d\lambda') = 0_n.$$

The above definition relates the function $c(\cdot)$ with a measure $\mu_\xi(d\lambda, d\lambda')$. No matter what $c(\cdot)$ is, if $\mu_\xi(d\lambda, d\lambda')$ is Lebesgue measure (on either $T \times T$ or the diagonal of $T \times T$) then the moving average representation is virile.

On the other hand, if $c(\cdot)$ has an absolutely convergent Fourier series,[2] so that

$$\sum_{j\in\mathbf{Z}} |\hat{c}_{kl}(j)| < \infty \qquad 1 \leqslant k \leqslant n,\ 1 \leqslant l \leqslant m,$$

then the moving average representation is virile no matter what $\mu_\xi(d\lambda, d\lambda')$ is. Between these two extremes, virility depends on both $\hat{c}(\cdot)$ and $\mu_\xi(d\lambda, d\lambda')$.

LEMMA 2.19. *Let $X_t = \sum_{j\in\mathbf{Z}} \hat{c}(j-t)\,\xi_j$ be an orthogonal moving average representation of X_t. If $\sup_{j\in\mathbf{Z}} \|\xi_j\| < \infty$, then it is an orthogonal weakly harmonizable virile moving average representation and X_t is a weakly harmonizable process.*

In [7], strongly harmonizable virile moving average representations are identified with particular covariance function representations. The latter are also termed virile.

LEMMA 2.20. *If an n-dimensional random process, X_t, has a stationary/strongly harmonizable/weakly harmonizable virile moving average representation* (2.3), *then it is a stationary/strongly harmonizable/weakly harmonizable process with virile covariance representation. Conversely, if* (2.3) *is a virile moving average representation, $c(\cdot)$ has full rank, and X_t is a strongly harmonizable* (*stationary*) *process then the moving average representation is strongly harmonizable* (*stationary*).

3. Linear Extrapolation

At time t_0, suppose we can observe the process, X_t for $t \leqslant t_0$, but not X_t in the future. We are required to predict X_t at time $t = t_0 + \tau$ for $\tau \in \mathbf{Z}^+$. Our prediction is based on what we know until time t_0. Thus each component of our prediction for $X_{t_0+\tau}$ will be an element of $H_X^-(t_0)$.

3.1. *Prediction*

DEFINITION 3.1. Given an n-dimensional random process, X_t, and $\tau \in \mathbf{Z}^+$, let $\hat{X}(t, \tau)$ be the element in $[H_X^-(t)]^n$ that best approximates $X_{t+\tau}$ in the $L^2(P)$ sense.

[2] Of interest is the following theorem:

THEOREM 2.18 (Zygmund). *Let $c(\cdot)$ be of bounded variation on T and assume $c(\cdot) \in \mathrm{Lip}_\alpha(T)$ for some $\alpha > 0$. Then $c(\cdot)$ has absolutely convergent Fourier series.* (see [6, Sect. 1.6].)

Ordinary Hilbert space theory shows that given the space $[H_X^-(\infty)]^n$, an element $X_{t+\tau}$ and a subspace $[H_X^-(t)]^n$, there exists a unique element $\hat{X}(t, \tau) \in [H_X^-(t)]^n$ that best approximates $X_{t+\tau}$ in norm. In other words, the above definition makes sense. One sees that $\hat{X}(t, \tau)$ is just the orthogonal projection of $X_{t+\tau}$ onto $[H_X^-(t)]^n$, with the error vector, $X_{t+\tau} - \hat{X}(t, \tau)$, being the perpendicular from $X_{t+\tau}$ to the space $[H_X^-(t)]^n$.

DEFINITION 3.2. A random process, X_t, is *deterministic* (or *linearly singular*) iff $H_X^-(-\infty) = H_X^-(\infty)$, and otherwise it is *nondeterministic*. It is *purely nondeterministic* (or *linearly regular*) iff $H_X^-(-\infty) = \{0\}$.

If X_t and Y_t have the same covariance function, the map $X_t \mapsto Y_t$ induces a Hilbert space isomorphism between $H_X^-(\infty)$ and $H_Y^-(\infty)$. Thus the question of whether a process is deterministic, non-deterministic, or purely nondeterministic can be answered by looking at its covariance function.

DEFINITION 3.3. The *innovation spaces*, $\{D_X(t)\}_{t \in Z}$, of a discrete random process, X_t, are defined as: $D_X(t) \stackrel{\text{def}}{=} H_X^-(t) \ominus H_X^-(t-1)$ (the orthogonal complement of $H_X^-(t-1)$ in $H_X^-(t)$).

A process, X_t, is purely nondeterministic iff $H_X^-(\infty)$ is the sum of its innovation spaces, i.e., if X_t has no "infinite past."

The following theorem was first discovered by H. Wold in the late 1930's for the discrete stationary case and later generalized to the theorem below essentially by H. Cramér [2].

THEOREM 3.4 (Wold's Decomposition Theorem). *For an n-dimensional random process, $X_t \in [L_0^2(P)]^n$, there exists a unique decomposition*

$$X_t = R_t + S_t,$$

where $E(R_s S_t^*) = 0_n$ for all $s, t \in Z$, and furthermore R_t is purely nondeterministic and S_t is deterministic. If X_t is weakly harmonizable, R_t and S_t are weakly harmonizable too.

Proof. Given a random process, X_t, if $\tilde{\pi}: H_X^-(\infty) \to H_X^-(-\infty)$ is the orthogonal projection of $H_X^-(\infty)$ onto $H_X^-(-\infty)$, then define

$$\pi: [H_X^-(\infty)]^n \to [H_X^-(-\infty)]^n$$

by $\pi \stackrel{\text{def}}{=} [\tilde{\pi}]^n$. Let $S_t \stackrel{\text{def}}{=} \pi X_t$ and $R_t \stackrel{\text{def}}{=} X_t - S_t$.

If X_t is weakly harmonizable, then X_t has a representation,

$$X_t = \int_T e^{it\lambda} Z(d\lambda).$$

Thus $S_t \doteq \int_T e^{it\lambda} \pi(Z(d\lambda))$, so S_t is weakly harmonizable too. Since X_t and S_t are weakly harmonizable, so is R_t and one has a weakly harmonizable decomposition. ∎

If a random process, X_t, is strongly harmonizable, one might ask if each member of its Wold's Decomposition is strongly harmonizable too. Lemma 2.14 supplies a sufficient condition, namely, if

$$\min\{\dim H_R^-(\infty), \dim H_S^-(\infty)\} < \infty,$$

then X_t has a strongly harmonizable Wold's Decomposition. A necessary and sufficient condition is not known at this time.

Although no algorithm exists for finding Wold's Decomposition, one frequently decomposes the linear extrapolation problem for a nondeterministic random process into two problems by trying to predict the future for the purely nondeterministic and deterministic components. Again, though no general algorithm exists, it is "theoretically possible"[3] to predict the future values of a deterministic process with certainty so the prediction problem in the deterministic case can be "declared solved." We will thus limit our study of linear prediction to purely nondeterministic processes.

3.2. *Filtered Processes*

To clarify the difference between forecasting for harmonizable and stationary processes, consider filtered processes.

DEFINITION 3.5. Given an n-dimensional harmonizable process,

$$X_t = \int_T e^{it\lambda} Z_X(d\lambda),$$

a p-dimensional harmonizable process, Y_t, is a *filtered process* with respect to X_t iff there exists a $\phi_Y(\cdot) \in L^2(F_X, p)$ such that $Y_t = \int_T e^{it\lambda} \phi_Y(\lambda)\, Z_X(d\lambda)$. The function $\phi_Y(\cdot)$ is called the *spectral characteristic* of the filtered process Y_t.

Suppose two p-dimensional processes, X_t and Y_t, are both filtered processes with respect to a third n-dimensional process, W_t. Furthermore, suppose X_t and Y_t are known from time t_0 until time t_1. The two processes

[3] In the stationary case, given $\tau > 0$, one can approximate X_t by $\sum_{j=1}^k a_j X_{t_j}$, where $t_j < t - \tau$ if X_t is deterministic. Then $\hat{X}(t, \tau) = X_{t+\tau}$ is approximately $\sum_{j=1}^k a_j X_{t_j+\tau}$ since there exists a one parameter family of unitary operators, U_t, such that $U_s X_t = X_{s+t}$. However, in the harmonizable case, a one parameter family of operators need not exist. Lacking such a family, "theoretically possible" only means that one is aware of the space of observables for future values of the process as well.

are said to have parallel observations iff $X_t = Y_t$ for $t \in [t_0, t_1]$. If W_t is known and is a stationary process then knowing X_t at even one point in time, t_0, is enough to determine X_t for all time.[4] Therefore, if W_t is a stationary process then parallel observations for X_t and Y_t (even if $[t_0, t_1]$ consists of just the time t_0) implies equality of the processes. The same cannot be said for harmonizable processes, as the following example illustrates.

EXAMPLE 3.6. Let $\tilde{W}_t$ be an n-dimensional stationary process with $r_{\tilde{W}}(s, t) = \delta_0(s-t)\, I_n$ and let π be the orthogonal projection of $H_{\tilde{W}}(\infty)$ onto $\overline{\mathrm{sp}}\{\tilde{W}_t^k : 1 \leqslant k \leqslant n,\ |t| > 4\}$. Let $W_t \overset{\text{def}}{=} [\pi]^n\, \tilde{W}_t$. Define X_t and Y_t to be the strongly harmonizable filtered processes with respect to W_t corresponding to the spectral characteristics I_n and $e^{i(\cdot)} I_n$. By Lemma 2.14, W_t, X_t, and Y_t are strongly harmonizable. One has $X_{-1} = X_0 = X_1 = Y_{-1} = Y_0 = Y_1 = 0$, yet the process X_t need not be the same as Y_t for all $t \in \mathbf{Z}$.

One is thus led to the following question: Given that X_t is a filtered process with respect to a strongly harmonizable process, W_t, for what $A \subseteq \mathbf{Z}$ are observations of X_t on A sufficient to determine the process X_t (or equivalently, its spectral characteristic $\phi_X(\cdot)$)? This problem is still unsolved.

4. Fundamental Moving Averages

4.1. *Introduction*

DEFINITION 4.1. Given a discrete purely nondeterministic random process, X_t, with a bounded covariance function, a *fundamental moving average representation* of X_t is a one-sided moving average,

$$X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j, \tag{4.4}$$

such that $\mathrm{sp}\{\xi_j^k : 1 \leqslant k \leqslant m\} = D_X(j)$ for each j. A *strongly harmonizable (stationary) fundamental moving average* is strongly harmonizable (stationary) moving average that is fundamental.

Only purely nondeterministic processes can have fundamental moving averages. All fundamental moving averages are one-sided orthogonal moving averages though the converse need not be true. In fact, if $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j$ is a one-sided orthogonal moving average representation

[4] Let $U_t : H_{\bar{W}}(\infty) \to H_{\bar{W}}(\infty)$ be the one parameter family of unitary operators such that $[U_t]^n W_0 = W_t$. Then $X_t = [U_{t-t_0}]^p X_{t_0}$.

of X_t, it follows that $H_X^-(t) \subseteq H_\xi^-(t)$ with equality iff the orthogonal moving average representation is fundamental.

Fundamental moving averages are useful in linear prediction theory since if (4.4) is such a representation of X_t, then for every $\tau \in \mathbf{Z}^+$

$$\hat{X}(t, \tau) = \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau)\, \xi_j. \tag{4.5}$$

The reason why "weakly harmonizable fundamental moving average" was not defined is because all fundamental moving averages are weakly harmonizable, as the following lemma points out.

LEMMA 4.2. *Every fundamental moving average representation,* (4.4), *of an n-dimensional random process, X_t, is a weakly harmonizable virile moving average. Furthermore, X_t is weakly harmonizable and rank $\hat{c}(0) = m$. If* (4.4) *is a strongly harmonizable* (*stationary*) *moving average then X_t and $\hat{X}(t, \tau)$ are strongly harmonizable* (*stationary*) *processes for each $\tau \in \mathbf{Z}^+$.*

Proof. Since a fundamental moving average is an orthogonal one, Lemma 2.19 shows that the fundamental moving average representation is virile and weakly harmonizable, and X_t is weakly harmonizable.

Note that if (4.4) is a fundamental moving average representation, $\hat{c}(j-t)\, \xi_j = \hat{X}(j, t-j) - \hat{X}(j, t-j-1)$. Thus $\hat{c}(j-t)\, \xi_j$ is just the orthogonal projection of X_t onto $[D_X(j)]^n$ and $D_X(j) = \mathrm{sp}\{(\hat{c}(0)\, \xi_j)^{(k)} : 1 \leqslant k \leqslant n\}$, which implies $\hat{c}(0)$ has full rank m.

Furthermore if (4.4) is a strongly harmonizable (stationary) moving average then Lemma 2.20 implies that X_t is a strongly harmonizable (stationary) process.

Since the covariance function of X_t is bounded then so is the covariance function of $\hat{X}(t, \tau)$ since X_t can be written as the orthogonal sum, $X_t = \hat{X}(t, \tau) + \sum_{j=t+1}^{t+\tau} \hat{c}(j-t-\tau)\, \xi_j$. Clearly if (4.4) is a strongly harmonizable (stationary) process then (4.5) is. Thus if (4.4) is a strongly harmonizable (stationary) moving average then Lemma 2.20 implies that $\hat{X}(t, \tau)$ is a strongly harmonizable (stationary) process. ∎

Since the innovation spaces are orthogonal, $r_\xi(s, t) = \delta_0(t-s)\, d(s)\, I_m$ in (4.4), where $d \colon \mathbf{Z} \to [0, \infty)$. Since ξ_j is an orthogonal random process, the boundedness of its covariance function is equivalent to the boundedness of the covariance of X_t, hence $d(\cdot)$ is a bounded function and

$$\begin{aligned} X_t &= \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j \\ &= \sum_{j=-\infty}^{t} \left[\left(\sup_{k \in \mathbf{Z}} \sqrt{d(k)} \right) \hat{c}(j-t) \right] \frac{\xi_j}{\sup_{k \in \mathbf{Z}} \sqrt{d(k)}}. \end{aligned}$$

Thus it can be assumed that $d\colon \mathbf{Z} \to [0, 1]$, i.e., $\sup_{s,t\in \mathbf{Z}} r_\xi(s, t) = 1$, without loss of generality.

DEFINITION 4.3. An n-dimensional random process, X_t, is *m-staggered* iff for each $k \in \mathbf{Z}$, either $\dim D_X(k) = m$ or $\dim D_X(k) = 0$. Define $J_X \stackrel{\text{def}}{=} \{k \in \mathbf{Z} : \dim D_X(k) \neq 0\}$.

If, as in Definition 4.1, X_t has an n-dimensional one-sided moving average representation (4.4) ($c(\cdot)$ is an $n \times m$ matrix valued function) then for each $j \in \mathbf{Z}$ one observes that $\dim \operatorname{sp}\{\xi_j^{(1)}, ..., \xi_j^{(m)}\}$ is either m or 0 since $r_\xi(s, t) = \rho(s, t)\, I_m$. Thus only staggered processes can have fundamental moving averages.

Let $\pi_k\colon H_X^-(\infty) \to D_X(k)$ be the orthogonal projection onto the kth innovation space, $D_X(k)$.

DEFINITION 4.4. An n-dimensional random process, X_t, is *aligned* iff it is m-staggered and there exist an $n \times m$ constant matrix, c, and an m-dimensional random process, ξ_k, such that

$$c\xi_k = [\pi_k]^n X_k. \tag{4.6}$$

Every n-dimensional, n-staggered process is aligned since one can let $c = I_n$ and $\xi_k = [\pi_k]^n X_k$. Simple examples show that not every n-dimensional, m-staggered process is aligned for $m < n$.

Given an n-dimensional m-staggered random process, X_t, let $\{\xi_j^{(1)}, ..., \xi_j^{(m)}\}$ be a basis for $D_X(j)$ (with $\{\xi_j^{(1)}, ..., \xi_j^{(m)}\} = \{0, ..., 0\}$ if $\dim D_X(j) = 0$) such that $E(\xi_j \xi_j^*) = K_j I_m$, where $K_j \geqslant 0$. For $j \in J_X$ there exists a unique $n \times m$ matrix, $c_j(0)$, of rank m such that $[\pi_j]^n X_j = c_j(0)\, \xi_j$. Similarly, for each $j \in J_X$ and $k \geqslant 0$, there exists an $n \times m$ matrix, $c_j(-k)$, (not necessarily of rank m) so that

$$[\pi_j]^n X_{j+k} = c_j(-k)\, \xi_j. \tag{4.7}$$

Every $n \times m$ matrix, A, of rank m, has a left inverse defined on $A(C^m)$. We will use the generalized inverse, $A^\dagger$ of A to represent this left inverse (see [8]).

DEFINITION 4.5. Given an m-staggered random process, X_t, with $j \in J_X$ and $k \geqslant 0$, the *innovation ratio*, $I(j, k)\colon c_j(0)(C^m) \to C^n$, is given by

$$I(j, k) \stackrel{\text{def}}{=} c_j(-k)\, c_j^\dagger(0). \tag{4.8}$$

One needs to show that the above definition of innovation ratio is independent of the original choice of basis elements for $D_X(j)$. To see this, let

$\{\eta_j^{(1)}, ..., \eta_j^{(m)}\}$ be another basis of $D_X(j)$ such that $E(\eta_j \eta_j^*) = k_j I_m$ where $k_j \geqslant 0$. There exists an $m \times m$ constant multiple of a unitary matrix, V_j, of rank m, so that $\xi_j = V_j \eta_j$. One now observes that $[\pi_j]'' X_j = c_j(0) V_j \eta_j$ and $[\pi_j]'' X_{j+k} = c_j(-k) V_j \eta_j$. Thus

$$I(j, k) = c_j(-k) c_j^\dagger(0) = [c_j(-k) V_j][V_j^{-1} c_j^\dagger(0)].$$

DEFINITION 4.6. An m-staggered random process, X_t, has *invariant innovation ratios* iff $I(j, k)$ is independent of $j \in J_X$, i.e., it is a function, $I(k)$, of only k.

LEMMA 4.7. *If a random process, X_t, has a fundamental moving average representation then it is an aligned process with invariant innovation ratios.*

Proof. Let $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t) \xi_j$ be a fundamental moving average representation for X_t. Observe that, $\hat{c}(0) \xi_j = [\pi_j]'' X_j$ for all $j \in \mathbf{Z}$ and for $j \in J_X$,

$$I(j, k) = \hat{c}(-k) \hat{c}^\dagger(0),$$

which is independent of j. ▌

One might now conjecture that the existence of invariant innovation ratios for a purely nondeterministic aligned process with a bounded covariance function would imply the existence of a fundamental moving average. While these are certainly necessary conditions, they are not sufficient, as the following construction shows.

Assume X_t is a purely nondeterministic aligned random process with invariant innovation ratios and a bounded covariance function. Let $\tilde{\xi}_k$ and c be as ξ_k and c are in Definition 4.4. Fix $j \in J_X$ and let V be an $m \times m$ matrix such that $E([V\tilde{\xi}_j][V\tilde{\xi}_j]^*) = I_m$. Define $c(0) \stackrel{\text{def}}{=} cV^{-1}$ and $\xi_k \stackrel{\text{def}}{=} V\tilde{\xi}_k$ for $k \in \mathbf{Z}$. Definition 4.4 is now valid for "$c = c(0)$" and ξ_k. For each $k \in Z$ let $c(-k) \stackrel{\text{def}}{=} I(k) c(0)$. Now using the notation of (4.7) and (4.8), note that $c_k(0) = c(0)$ by definition, and for each $k \in \mathbf{Z}$ and $t \geqslant k$ we have

$$\begin{aligned} [\pi_k]'' X_t &= c_k(k-t) \xi_k && \text{by (4.7)} \\ &= [c_k(k-t)][c_k^\dagger(0) c_k(0) \xi_k] \\ &= [I(t-k)][[\pi_k]'' X_k] && \text{by (4.7)} \\ &= [c(k-t) c^\dagger(0)][c(0) \xi_k] && \text{by (4.6)} \\ &= c(k-t) \xi_k. \end{aligned}$$

One now has

$$X_t = \sum_{k=-\infty}^{t} [\pi_k]^n X_t = \sum_{k=-\infty}^{t} c(k-t)\,\xi_k. \tag{4.9}$$

However one cannot conclude that (4.9) is a moving average representation since

1. the function $c(\cdot)$ need not be a Fourier transform (since $1/\|\xi_j\|$, for $j \in J_X$, need not be bounded) and
2. for some $k \in \mathbf{Z}$ it is possible that $E(\xi_k \xi_k^*) \neq K_k I_m$ for any $K_k \geqslant 0$.

If it is known ahead of time that X_t has a (strongly harmonizable/stationary) fundamental moving average, then the above construction will give such a (strongly harmonizable/stationary) fundamental moving average.

LEMMA 4.8. *Let X_t have two fundamental moving average representations,*

$$X_t = \sum_{k=-\infty}^{t} \hat{c}(k-t)\,\xi_k = \sum_{k=-\infty}^{t} \hat{a}(k-t)\,\eta_k.$$

Then there is a $K > 0$ and an $m \times m$ unitary matrix, V, such that $\xi_k = KV\eta_k$ and $a(k) = Kc(k)\,V$ for all $k \in \mathbf{Z}$.

Proof. Fixing $j \in J_X$, there exists a $K > 0$ and an $m \times m$ unitary matrix, V, such that $\xi_j = KV\eta_j$. Since

$$[\pi_j]^n X_j = \hat{c}(0)\,\xi_j = \hat{c}(0)\,KV\eta_j = \hat{a}(0)\,\eta_j,$$

and $\hat{c}(0)$ and $\hat{a}(0)$ have rank m, one can conclude that $\hat{a}(0) = K\hat{c}(0)\,V$. Now for each $k \in Z$

$$\hat{a}(-k) = I(k)\,\hat{a}(0) = I(k)\,\hat{c}(0)\,KV = \hat{c}(-k)\,KV$$

and

$$\begin{aligned}\xi_k &= \hat{c}^\dagger(0)[\pi_k]^n X_k = [\hat{a}(0)(KV)^{-1}]^\dagger\,[\pi_k]^n X_k \\ &= (KV)\,\hat{a}^\dagger(0)[\pi_k]^n X_k = KV\eta_k. \quad \blacksquare\end{aligned}$$

4.2. *Stationary Fundamental Moving Averages*

A stationary fundamental moving average, $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$, is, up to a multiplicative constant, an orthonormal moving average. It will be assumed that a stationary fundamental moving average is orthonormal, unless otherwise stated.

Pure nondeterminism and the existence of a stationary fundamental moving average are equivalent properties for stationary processes as the following theorem of Y. Rozanov shows (see [10, p. 56]). A short proof is included for completeness.

THEOREM 4.9. *An n-dimensional stationary process, X_t, has a stationary fundamental moving average iff it is purely nondeterministic.*

Proof. It is clear that the existence of a stationary fundamental moving average implies that X_t is purely nondeterministic. For the converse, let U_t: $H_X^-(\infty) \to H_X^-(\infty)$ be a one parameter family of unitary operators such that $[U_s]^n X_t = X_{s+t}$ and let ξ_0 be an m-dimensional random vector such that

1. $\mathrm{sp}\{\xi_0^{(1)}, ..., \xi_0^{(m)}\} = D_X(0)$ and
2. $E(\xi_0 \xi_0^*) = I_m$.

Define $\xi_t \stackrel{\text{def}}{=} [U_t]^m \xi_0$. It follows that $\mathrm{sp}\{\xi_t^{(1)}, ..., \xi_t^{(m)}\} = D_X(t)$ and that $E(\xi_s \xi_t^*) = \delta_0(t-s) I_m$. Let $a(\cdot)$ be an $n \times m$ matrix function on the non-positive integers defined such that $a(j)\,\xi_j$ is the orthogonal projection of X_0 onto the jth innovation space. Then $X_t = \sum_{j=-\infty}^{t} a(j-t)\,\xi_j$. Since the ξ_j are orthonormal, $a(\cdot) \in l^2$ (that is to say that $tr \sum_{j=-\infty}^{0} a(j)\,a^*(j) < \infty$) and therefore there exists a matrix valued function, $c(\cdot)$, on T such that $a(\cdot) = \hat{c}(\cdot)$. Thus $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ is a stationary fundamental moving average representation of X_t. ∎

A related result is the following (see [10, p. 64]):

THEOREM 4.10. *An n-dimensional stationary process of rank n is purely nondeterministic iff it has a spectral density, $f(\cdot)$ $(= dF_X(\lambda)/d\lambda)$, with respect to Lebesgue measure such that*

$$\int_T \log \det f(\lambda)\, d\lambda > -\infty.$$

The above theorem, along with Theorem 4.9, leads one to ask as to when an n-dimensional stationary process of rank $m \neq n$ might be purely nondeterministic.

DEFINITION 4.11. The function $c\colon T \to \mathcal{M}_{n,m}$ is *maximal* iff

1. $c(\cdot) \in H^2(T)$ (the Hardy space consisting of all functions in $L^2(T)$ such that their Poisson integrals are analytic on the unit disk, see [4, p. 39]) and,

2. if $a(\cdot) \in H^2(T)$ and $c(\lambda)\, c^*(\lambda) = a(\lambda)\, a^*(\lambda)$, then $c(0)\, c^*(0) \geqslant a(0)\, a^*(0)$[5] where $c(0)$ and $a(0)$ are the values at zero of the analytic continuation of $c(\cdot)$ and $a(\cdot)$ respectively to the unit disk in C.

Using a theorem of G. Szegö, Y. Rozanov [10] discusses how to determine the maximality of a matrix valued function on T. For instance, if $c(\cdot)$ is a scalar function, it is maximal iff it is an outer function. An outer function [4] is one whose analytic continuation to the unit disk can be expressed as

$$c(z) = \lambda \exp\left[\frac{1}{2\pi}\int_T \frac{e^{i\theta}+z}{e^{i\theta}-z}\, k(\theta)\, d\theta\right],$$

where $k(\cdot)$ is a real-valued integrable function on the circle and λ is a complex number of modulus 1. In the case that $c(\cdot)$ takes values in $\mathscr{M}_{n,n}$, $c(\cdot)$ is maximal iff the analytic continuation of $c(\cdot)$ to the unit disk has the property,

$$|\det c(0)|^2 = (2\pi)^n \exp\left\{\frac{1}{2\pi}\int_T \log\det\,[c(\lambda)\, c^*(\lambda)]\, d\lambda\right\}.$$

Several other cases are also discussed.

The reason why maximal functions are important is seen from (see [10, p. 60]):

THEOREM 4.12 (Rozanov). *An orthonormal moving average representation,* $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j$, *is a stationary fundamental moving average representation iff* $c(\cdot)$ *is maximal.*

The situation changes radically for harmonizable processes as will be shown in the next subsection.

4.3. *Harmonizable Fundamental Moving Averages*

H. Cramér [3] has shown that every covariance function of a strongly harmonizable process has a "standard form" from which one can ascertain the determinism properties. However, given a strongly harmonizable process, there is, as yet, no known method of expressing its covariance function in Cramér's "standard form."

For a purely nondeterministic harmonizable process, X_t, (unlike purely nondeterministic stationary processes) the dimension of the innovation spaces, $D_X(t)$, need not be the same for all $t \in \mathbf{Z}$ as the following example shows.

[5] That is, the difference $c(0)\, c^*(0) - a(0)\, a^*(0)$ is a positive semi-definite matrix.

EXAMPLE 4.13. Define $\{Y_j\}_{j\in\mathbf{Z}}$ to be orthonormal n-dimensional random variables. Then Y_t is a purely nondeterministic stationary process. Let $X_t = Y_t$ for $t \neq 0$ and $X_0 = 0$. Lemma 2.14 shows that X_t is strongly harmonizable. However the dimension of $D_X(0)$ is zero, while $D_X(t) = D_Y(t)$ for $t \neq 0$.

PROPOSITION 4.14. *Every purely nondeterministic harmonizable n-dimensional process, X_t, has a purely nondeterministic stationary dilation.*

Proof. Let (Y_t, π) be a stationary dilation of X_t via Theorem 2.12. By Wold's decomposition $Y_t = R_t + S_t$ where R_t is purely nondeterministic and S_t is deterministic. Since

$$\pi[H_S^-(\infty)]^n = \pi[H_S^-(-\infty)]^n \subseteq \pi[H_Y^-(-\infty)]^n = [H_X^-(-\infty)]^n = \{0\},$$

it follows that $\pi[H_S^-(\infty)]^n = \{0\}$. Thus $X_t = \pi Y_t = \pi R_t$ and (R_t, π) is a purely nondeterministic stationary dilation of X_t. ∎

The following example shows that the converse of the above lemma is false, i.e., that the orthogonal projection of a purely nondeterministic stationary processes need not be purely nondeterministic.

EXAMPLE 4.15. Define $\{Y_j\}_{j\in\mathbf{Z}}$ to be orthonormal n-dimensional random variables. Then Y_t is a purely nondeterministic stationary process. Define $W \stackrel{\text{def}}{=} Y_0 + \sum_{j\neq 0} Y_j/j^2$ and let $\pi(\cdot)$ be the orthogonal projection onto the space spanned by W. Lemma 2.14 shows that $X_t \stackrel{\text{def}}{=} \pi Y_t$ is a strongly harmonizable process. However, X_t is not purely nondeterministic since $H_X^-(-\infty)$ is the space spanned by $\{W^{(1)}, ..., W^{(n)}\}$.

The next example shows that the analog of Theorem 4.10 for harmonizable processes is again false.

EXAMPLE 4.16. Let Y be a one dimensional random variable with $E(Y\bar{Y}) = 1$. Also let $c(\cdot) \in H^2(T)$ be a nonzero function such that $\max\{j \in \mathbf{Z} : \hat{c}(j) \neq 0\} = \infty$. Letting $X_t \stackrel{\text{def}}{=} \hat{c}(-t)\, Y$, the covariance function of X_t is

$$\begin{aligned} r_X(s,t) &= E((\hat{c}(-s)\, Y)\overline{(\hat{c}(-t)\, Y)}) \\ &= \hat{c}(-s)\,\overline{\hat{c}(-t)} = \iint_{T\times T} e^{is\lambda - it\lambda'} c(\lambda)\,\overline{c(\lambda')}\, d\lambda\, d\lambda'. \end{aligned}$$

Thus the spectral density function of X_t (with respect to Lebesgue measure) is $f_X(\lambda, \lambda') = c(\lambda)\,\overline{c(\lambda')}$. Furthermore,

$$\iint_{T\times T} \log |c(\lambda)\,\bar{c}(\lambda')|\,d\lambda\,d\lambda'$$

$$= 2\int_T \log |c(\lambda)|\,d\lambda = \int_T \log |c(\lambda)\,\bar{c}(\lambda)|\,d\lambda > -\infty$$

(see [4, p. 53] for the inequality). However, it is clear that

$$H_X^-(-\infty) = H_X^-(\infty) = \{aY : a \in C\}$$

so that X_t is not purely nondeterministic (indeed it is deterministic).

DEFINITION 4.17. A purely nondeterministic weakly harmonizable process, X_t, has a *fundamental stationary dilation*, $(Y_t, \pi, c(\cdot))$, iff

1. (Y_t, π) is a stationary dilation of X_t and
2. Y_t has a stationary fundamental moving average representation,

$$Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\tilde{\xi}_j$$

such that $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\pi\tilde{\xi}_j$ is an orthogonal moving average representation of X_t.

Notice that in the above definition, $\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\pi\tilde{\xi}_j$ must be a weakly harmonizable virile moving average by Lemma 2.19.

EXAMPLE 4.18. Let $Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\tilde{\xi}_j$ be a one dimensional fundamental moving average representation of a stationary process. Let $H_X^-(\infty) = \overline{\mathrm{sp}}\{\tilde{\xi}_j : j \neq 0\}$ and define the orthogonal projection $\pi: H_Y^-(\infty) \to H_X^-(\infty)$ by

$$\xi_j \overset{\text{def}}{=} \pi\tilde{\xi}_j = \begin{cases} \tilde{\xi}_j & \text{if } j \neq 0, \\ 0 & \text{if } j = 0 \end{cases}.$$

One now concludes that $(Y_t, \pi, c(\cdot))$ is a fundamental stationary dilation of $X_t \overset{\text{def}}{=} \pi Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$. Lemma 2.14 implies that both ξ_j and X_j are strongly harmonizable processes.

One might conjecture that every aligned purely nondeterministic harmonizable process has a fundamental stationary dilation. This is not so, as the next example demonstrates.

EXAMPLE 4.19. Let Y_t be as in Example 4.18 and let π' be the orthogonal projection such that

$$\pi'\tilde{\xi}_j = \begin{cases} \tilde{\xi}_j & \text{if } j \notin \{1, 2\}, \\ \dfrac{\sqrt{2}}{2}(\tilde{\xi}_1 + \tilde{\xi}_2) & \text{if } j \in \{1, 2\} \end{cases}.$$

Letting $X_t \stackrel{\text{def}}{=} \pi' Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \pi'\tilde{\xi}_j$, one notes that $D_X(1) = \text{sp}(\tilde{\xi}_1 + \tilde{\xi}_2)$. Thus X_t does not have invariant innovation ratios since $I(1, k) = [\hat{c}(-k) + \hat{c}(1-k)]\, \hat{c}^{\dagger}(0)$, while $I(3, k) = \hat{c}(-k)\, \hat{c}^{\dagger}(0)$. It follows that X_t does not have a fundamental moving average representation (by Lemma 4.7), so it cannot have a fundamental stationary dilation.

THEOREM 4.20. *An n-dimensional X_t, has a fundamental stationary dilation $(Y_t, \pi, c(\cdot))$ iff it has a fundamental moving average representation, $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j$, with $c(\cdot)$ a maximal function. Furthermore, the fundamental moving average of X_t is strongly harmonizable iff X_t is.*

Proof. ($\Rightarrow$) Let

$$Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \tilde{\xi}_j$$

be a stationary fundamental representation of Y_t. Theorem 4.12 shows that $c(\cdot)$ is maximal so it suffices to show that the moving average representation,

$$X_t = \pi Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \pi\tilde{\xi}_j = \sum_{j=-\infty}^{t} \hat{c}(j-t)\, \xi_j, \tag{4.10}$$

is fundamental.

Since $H_Y^-(j) = H_{\tilde{\xi}}^-(j)$, it follows that $H_X^-(j) = \pi H_Y^-(j) = \pi H_{\tilde{\xi}}^-(j)$. Thus

$$\begin{aligned} H_X^-(j-1) \oplus D_X(j) = H_X^-(j) & \\ &= \pi H_Y^-(j) \\ &= \pi(H_Y^-(j-1) \oplus D_Y(j)) \\ &= H_X^-(j-1) \oplus \pi D_Y(j). \end{aligned}$$

Now since $H_X^-(j-1) \perp D_X(j)$ and the definition of $(Y_t, \pi, c(\cdot))$ shows that $\pi H_Y^-(j-1)=H_X^-(j-1)$ is perpendicular to $\pi D_Y(j)$ it follows that for all j,

$$D_X(j)=\pi D_Y(j)=\pi \operatorname{sp}\{\tilde{\xi}_j^k : 1 \leqslant k \leqslant m\}=\operatorname{sp}\{\xi_j^k : 1 \leqslant k \leqslant m\}.$$

Thus $\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ is a fundamental representation of X_t.

($\Leftarrow$) Let $X_t=\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ be a fundamental moving average representation of X_t with $r_\xi(s,t)=\rho(s)\,\delta_0(t-s)\,I_m$ and $\sup_{s\in\mathbf{Z}} p(s)=1$. Let $\{\eta\}_{t\in\mathbf{Z}}$ be a collection of m-dimensional random variables such that

1. $E(\eta_s\eta_t^*)=\delta_0(t-s)\,I_m$ and
2. $E(\eta_s\xi_t^*)=0_m$ for each s, $t\in\mathbf{Z}$.

It may be necessary to augment (Ω, Σ, P) to a larger probability space to find the above η_t's. Let $\tilde{\xi}_t \overset{\text{def}}{=} \xi_t+\sqrt{1-\rho(t)}\,\eta_t$, let $Y_t \overset{\text{def}}{=} \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\tilde{\xi}_j$ and let π be the orthogonal projection of $H_Y^-(\infty)$ onto $H_X^-(\infty)$. Since $c(\cdot)$ is a maximal function, Theorem 4.12 implies that $Y_t=\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\tilde{\xi}_j$ is a stationary fundamental moving average representation, and thus $(Y_t, \pi, c(\cdot))$ is a fundamental stationary dilation of X_t.

As for the final assertion, the spectral characteristic, $\phi_\xi(\cdot)$ of $\tilde{\xi}_t$ with respect to Y_t exists since $H_{\tilde{\xi}}^-(\infty)=H_Y^-(\infty)$ and Y_t and ξ_t are stationarily correlated [10, p. 59]. Then $\tilde{\xi}_j=\int_T e^{ij\lambda}\phi_\xi(\lambda)\,Z_Y(d\lambda)$ and

$$\xi_j=\int_T e^{ij\lambda}\phi_\xi(\lambda)\,\pi Z_Y(d\lambda)=\int_T e^{ij\lambda}\phi_\xi(\lambda)\,Z_X(d\lambda).$$

Thus

$$E(\xi_s\xi_t^*)=\iint_{T\times T} e^{is\lambda-it\lambda'}\,\phi_\xi(\lambda)\,F_X(d\lambda, d\lambda')\,\phi_\xi^*(\lambda').$$

It now follows that ξ_j is strongly harmonizable if X_t is. Lemmas 4.2 and 2.20 reveal that X_t is strongly harmonizable if ξ_j is. Thus X_t is strongly harmonizable iff its fundamental moving average representation (4.10) is. ∎

COROLLARY 4.21. *Let $X_t=\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ be an orthogonal moving average representation of an n-dimensional weakly harmonizable process, X_t, and let $c(\cdot)$ be a maximal function. Then there exists a fundamental stationary dilation of X_t and $X_t=\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ is a fundamental moving average representation.*

Proof. The construction of a fundamental stationary dilation is given in the second part of the proof of Theorem 4.20. The fact, that $X_t=\sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ is a fundamental moving average representation now follows from Theorem 4.20. ∎

A question that Theorem 4.20 leaves unanswered is: Can a harmonizable process have a fundamental moving average representation yet not have a fundamental stationary dilation? Equivalent to an answer of "no" (see Lemma 4.8) is the existence of a fundamental moving average $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ where $c(\cdot)$ is not a maximal function.

A main result on prediction is included in the following:

THEOREM 4.22. *Let X_t be an n-dimensional strongly harmonizable process with a fundamental stationary dilation, $(Y_t, \pi, c(\cdot))$, and let $X_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j$ be a fundamental moving average representation of X_t of rank n. Fix $\tau \in Z^+$. Then $\hat{X}(t,\tau)$ is an n-dimensonal strongly harmonizable process with a fundamental stationary dilation, $(\hat{Y}(t,\tau), \pi, c(\cdot)\,e^{i\tau\cdot})$. There exists a $\phi_\tau(\cdot) \in L^2(F_X, n)$,*

$$\phi_\tau(\lambda) \stackrel{\text{def}}{=} e^{i\lambda\tau}\left[c(\lambda) - \sum_{j=-\tau+1}^{0} \hat{c}(j)\,e^{i\lambda j}\right] c^\dagger(\lambda),$$

where

$$c(\lambda) = \sum_{j=-\infty}^{0} \hat{c}(j)\,e^{i\lambda j}$$

such that

$$\hat{X}(t,\tau) = \int_T e^{it\lambda}\phi_\tau(\lambda)\,Z_X(d\lambda). \tag{4.11}$$

Furthermore

$$r_{\hat{X}(\cdot,\tau)}(s,t) = \iint_{T\times T} e^{i(s+\tau)\lambda - i(t+\tau)\lambda'} \times\left(\sum_{j=-\infty}^{-\tau} \hat{c}(j)\,e^{i\lambda j}\right)\left(\sum_{k=-\infty}^{-\tau} \hat{c}(k)\,e^{i\lambda' k}\right)^{*} \mu_\xi(d\lambda, d\lambda'), \tag{4.12}$$

and the error of prediction is

$$X_{t+\tau} - \hat{X}(t,\tau) = \int_T e^{i(t+\tau)\lambda}\left[I_n - \left(\sum_{j=-\infty}^{-\tau} \hat{c}(j)\,e^{i\lambda j}\right) c^\dagger(\lambda)\right] Z_X(d\lambda). \tag{4.13}$$

Finally

$$\begin{aligned} &\| X_{t+\tau} - \hat{X}(t,\tau)\|^2_{[L_0^2(P)]^n} \\ &= \iint_{T\times T} e^{i(t+\tau)\lambda - i(t+\tau)\lambda'} \times \operatorname{tr}\left[I_n - \left(\sum_{j=-\infty}^{-\tau} \hat{c}(j)\,e^{i\lambda j}\right) c^\dagger(\lambda)\right] \\ &\quad \times \left[I_n - \left(\sum_{k=-\infty}^{-\tau} \hat{c}(k)\,e^{i\lambda' k}\right) c^\dagger(\lambda')\right]^{*} \mu_\xi(d\lambda, d\lambda'). \end{aligned} \tag{4.14}$$

Proof. Let $Y_t = \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\tilde{\xi}_j$ be a stationary fundamental moving average representation of Y_t where $\pi\tilde{\xi}_j = \xi_j$. Then $\hat{Y}(t,\tau) = \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau)\,\tilde{\xi}_j$ is a stationary fundamental moving average representation of $\hat{Y}(t,\tau)$. Letting $a_\tau(\cdot) = c(\cdot)\,e^{i\tau(\cdot)}$, one can write $\hat{Y}(t,\tau) = \sum_{j=-\infty}^{t} \hat{a}_\tau(j-t)\,\tilde{\xi}_j$. It is clear that $\hat{X}(t,\tau) = \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau)\;\pi\tilde{\xi}_j = \sum_{j=-\infty}^{t} \hat{a}_\tau(j-t)\,\xi_j$ is a fundamental moving average representation of $\hat{X}(t,\tau)$. Thus $(\hat{Y}(t,\tau), \pi, a_\tau(\cdot))$ is a fundamental stationary dilation of $\hat{X}(t,\tau)$. Since $\hat{X}(t,\tau)$ is strongly harmonizable, Theorem 4.20 reveals that $\sum_{j=-\infty}^{t} \hat{a}_\tau(j-t)\,\xi_j$ is a strongly harmonizable fundamental moving average representation.

The virility of fundamental moving averages allows the interchange of summation and integral symbols in the equations below.

One has

$$\begin{aligned}
\int_T e^{it\lambda} Z_X(d\lambda) = X(t) &= \sum_{j=-\infty}^{t} \hat{c}(j-t)\,\xi_j = \sum_{j=-\infty}^{t} \hat{c}(j-t) \int_T e^{ij\lambda}\, Z_\xi(d\lambda)\\
&= \int_T e^{it\lambda} \left(\sum_{j=-\infty}^{t} \hat{c}(j-t)\, e^{i(j-t)\lambda} \right) Z_\xi(d\lambda)\\
&= \int_T e^{it\lambda} c(\lambda)\, Z_\xi(d\lambda).
\end{aligned}$$

Thus $Z_X(d\lambda) = c(\lambda)\,Z_\xi(d\lambda)$ or $Z_\xi(d\lambda) = c^{\dagger}(\lambda)\,Z_X(d\lambda)$. Now by (4.5) we have,

$$\begin{aligned}
\hat{X}(t,\tau) &= \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau)\,\xi_j = \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau) \int_T e^{ij\lambda} Z_\xi(d\lambda)\\
&= \int_T e^{i(t+\tau)\lambda} \sum_{j=-\infty}^{t} \hat{c}(j-t-\tau)\, e^{i(j-t-\tau)\lambda}\, Z_\xi(d\lambda)\\
&= \int_T e^{i(t+\tau)\lambda} \sum_{j=-\infty}^{-\tau} \hat{c}(j)\, e^{ij\lambda} Z_\xi(d\lambda)\\
&= \int_T e^{i(t+\tau)\lambda} \left[c(\lambda) - \sum_{j=-\tau+1}^{0} \hat{c}(j)\, e^{ij\lambda} \right] Z_\xi(d\lambda)\\
&= \int_T e^{it\lambda} \left[e^{i\tau\lambda} \left(c(\lambda) - \sum_{j=-\tau+1}^{0} \hat{c}(j)\, e^{ij\lambda} \right) c^{\dagger}(\lambda) \right] Z_X(d\lambda),
\end{aligned}$$

which is just (4.11).

The rest of the proof now follows directly from routine calculations. ∎

The above theorem can be proved with other, somewhat specialized, assumptions on $c(\cdot)$ besides maximal rank. These results can be obtained by observing (with the above notation) that if there exists a $\phi_\tau(\cdot)$ such that

$$\hat{Y}(t, \tau) = \int_T e^{it\lambda} \phi_\tau(\lambda)\, Z_Y(d\lambda),$$

then

$$\hat{X}(t, \tau) = \int_T e^{it\lambda} \phi_\tau(\lambda)\, \pi Z_Y(d\lambda) = \int_T e^{it\lambda} \phi_\tau(\lambda)\, Z_X(d\lambda).$$

Some necessary conditions on $c(\cdot)$ for the existence of the spectral characteristic, $\phi_\tau(\cdot)$, have been found by Y. Rozanov (see [10, Chap. 2]) for the stationary case; full maximal rank being one of these conditions.

Acknowledgments

The author thanks the referee for his careful reading of the submitted work and Professor M. M. Rao for his advice and encouragement. This research was partially supported by ONR Contract N00014–84–K–0356.

References

[1] Chang, D. and Rao, M. M. (1986). Bimeasures and nonstationary processes. In *Real and Stochastic Analysis* (M. M. Rao, Ed.), pp. 7-118. John Wiley, New York.
[2] Cramér, H. Décompositions orthoganales de certains processus stochastiques. *Ann. Fac. Sci. Univ. Clermont* **11** 15–21.
[3] Cramér, H. (1959). On the linear prediction problem for certain stochastic processes. *Ark. Mat.* **4** (6) 45–53.
[4] Hoffman, K. (1962). *Banach Spaces of Analytic Functions.* Prentice–Hall, New Jersey.
[5] Kakihara, Y. A note on harmonizable and V-bounded processes. *J. Multivariate Anal.* **16** (1) 140–156.
[6] Katznelson, Y. (1976). *An Introduction to Harmonic Analysis.* Dover, New York.
[7] Mehlman, M. H. (1991). Structure and moving average representation for multidimensional strongly harmonizable processes. *Stochastic Anal. Appl.* **9** (3) 323–361.
[8] Penrose, R. A. (1955). A generalized inverse for matrices. *Proc. Cambridge Philos. Soc.* **51** 406–413.
[9] Rao, M. M. (1992). Harmonizable processes: Structure theory. *L'Enseignement Math.* **28** (2) 295–351.
[10] Rozanov, Yu. A. (1967). *Stationary Random Processes.* Holden–Day, San Francisco.